Legal Person- or Agenthood of Artificial Intelligence Technologies

Tanel Kerikmäe
Peeter Müürsepp[1]
Henri Mart Pihl

Department of Law
Tallinn University of Technology
Akadeemia tee 3
Tallinn 12618, Estonia
Email: tanel.kerikmae@taltech.ee
Email: peeter.muursepp@taltech.ee
Email: pihl.henri@gmail.com

Ondrej Hamuľák[2]
Hovsep Kocharyan[3]

Faculty of Law
Palacký University Olomouc
tr. 17. listopadu 8
Olomouc 771 11, Czech Republic
Email: ondrej.hamulak@upol.cz
Email: hovsep.kocharyan01@upol.cz

Abstract: Artificial intelligence (AI) is developing rapidly. There are technologies available that fulfil several tasks better than humans can and even behave like humans to some extent. Thus, the situation prompts the question whether AI should be granted legal person- and/or agenthood? There have been similar situations in history where the legal status of slaves or indigenous peoples was discussed. Still, in those historical questions, the subjects under study were always natural persons, i.e., they were living beings belonging to the species Homo sapiens. We analyse the situation

Tanel Kerikmäe, Peeter Müürsepp,
Ondrej Hamuľák, Hovsep Kocharyan, Henri Mart Pihl

from moral-ethical and practical perspectives. The final conclusion is that the currently existing AIs are still so far removed from humans that there is simply no need to think seriously about legal person- or agenthood. Doing so would mean imposing obligations on the AI to follow. This, in turn, would mean that certain rights in relation to those obligations would have to be granted as well. By all evidence, this is something that humans are not ready to do yet and might never get that far.

Keywords: *artificial intelligence, ethical and moral, legal agenthood, legal personhood, the missing something argument, practice run*

Introduction

Artificial intelligence (AI) is the central topic in many areas of knowledge and all possible aspects of it are under study. Our aim is to contribute to building the philosophical foundation of AI and on the basis of this to analyse the very basic legal aspect of AI—namely, whether legal person- or agenthood would apply to AI. In the article, we shall survey various arguments or subjects related to the theme before proceeding to explain our own position.

In a philosophical analysis we have to start with a definition or at least an explanation of what we are talking about when we speak about AI. It is not an easy task and cannot be conclusively resolved but we can present a kind of framed explanation that does the job for us for the time being. We have the European Commission High-Level Expert Group on AI definition from April 2019 at our disposal:

> Artificial intelligence (AI) systems are software (and possibly also hardware) systems that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge derived from this data and deciding the best action(s) to take to achieve the given goal. AI systems can also be designed to learn to adapt their behaviour by analysing how the environment is affected by their previous actions. (Sendler, 2019)

The definition looks intriguing as it assigns an AI tasks that human beings implement as well. The only clear difference is that we understand AIs as human-designed systems.

There is a shorter, more concise version of the definition provided by the EU Commission's Communication on AI: "Artificial intelligence (AI) refers to systems that display intelligent behaviour by analysing their environment and taking actions—with some degree of autonomy—to achieve specific goals" (Communication COM(2018) 237 final). In what follows, we shall adhere more directly to this version of the definition of AI. After all, similarity to human activity is well understood, as Frank Rosenblatt, one of the founding fathers of AI, used the human brain model to elaborate artificial perception already back in 1958 (Rosenblatt, 1958).

AI is certainly one of the central topics of the Fourth Industrial Revolution. It is interesting that the process is called an Industrial Revolution, although the most visible developments concern rather information communication technology. The use cases developed and currently under development by both public and private sector are making news almost every day. As always with relevant phenomena, there are visibly different societal expectations of and also fears related with, for example, using an AI as a data analyst against certain groups of people, stealing the customer's choice and workplaces. Also, biased or insufficient data can bring unwelcomed results or form a ground for decisions that are not serving the interests of society. Others are welcoming the triumph of AIs and are ready to accept a new type of species who may have a status comparable to humans. These developments raise concerns that promote initiatives for working out central directives for handling the ethical concerns connected to the AI. For instance, UNESCO has recently embarked on a two-year process to elaborate the first global standard-setting instrument on the ethics of artificial intelligence in the form of a Recommendation. The process was initiated following the decision of UNESCO's General Conference at its 40th session in November 2019 (UNESCO, n.d.). The first draft of the document was completed in April and May 2020 by UNESCO Ad Hoc Expert Group (AHEG) and was sent out to a wide variety of organisations and individuals from different walks of life for consultation.[4] The current draft text contains a very straightforward clause (that may still be changed during the consultation period): "When developing regulatory frameworks, governments should, in particular, take into account that responsibility and accountability must always lie with a natural or legal person; responsibility should not be delegated to an AI system, nor should a legal personality be given to an AI system" (IRCAI, 2020). In September 2020,

---

[4]   One of the authors, Prof. Tanel Kerikmäe has been a national consultant in this process. Dr. Peeter Müürsepp has been involved at the consultation stage.

Tanel Kerikmäe, Peeter Müürsepp,
Ondrej Hamuľák, Hovsep Kocharyan, Henri Mart Pihl

the preliminary report and first draft of the Recommendation were transmitted to the UNESCO Member States for their written comments by 31 December 2020. (UNESCO, n.d.)

This attempt to deny legal personality does not still solve all the problems raised in this article. The main question in this complicated and novel situation is: what is the difference between humans and the AI? If we fail to single out any major difference then why should we not accept an AI as equal to humans in any respect including legal person- and agenthood. Before taking any final decisions, however, we should always remember John Searle's Chinese room argument (Searle, 1980). Having the impression that an AI understands the situations does not necessarily mean that it really does. On the other hand, we humans also do not always understand what our fellow citizens believe we do. However, our reasons for not understanding are still different from those in the Chinese room thought experiment.

We have reached the era where technologies with at least some kind of personhood are available already. Let us take a look at some of them.

Examples of technologies with personhood

Despite all the controversies, the technologies that have personhood already exist and even if most of the legal scholars are not favouring this development with exhilaration, the issue whether to grant similar rights with humans to bots or algorithms cannot be ignored.

One of the most famous case is Sophia—an advanced human-like robot, developed by a Hong Kong based company Hanson Robotics. Created in 2016, it is the first robot to be granted citizenship by Saudi Arabia in 2017. It (or should we say "she") is also the Innovation Ambassador for the United Nations Development Programme and has been included in media as a public figure (TV shows such as the *Tonight Show* and *Good Morning Britain*). Sophia's design and primary purpose was to communicate with humans. She is able to read and use facial expressions, hold conversations on predetermined topics, walk and even draw (Hansen Robotics, n.d.). For additional reading on what Sophia is and what she could mean for intellectual property rights see Rocha, 2018.

Nadine is an advanced human-like robot developed by the Institute for Media Innovation in Singapore's Nanyang Technological University. Nadine is able to answer questions in different languages, remember the people she has interacted with and show emotions through expressions and gestures. She also has a "personality", meaning that her mood and answers can differ, depending on what you say to her and how you say it. Nadine is designed to be a companion robot, she can assist people with special needs, read stories, show images, set on Skype sessions, send emails, and communicate with the family members. She can play the role of a personal, private coach, always available when nobody is there. (IMI, n.d.) To read more about Nadine and AI in the medical field, see Zapusek, 2017.

VITAL, which stands for Validating Investment Tool for Advancing Life Sciences, unlike Nadine and Sophia, does not have a human-like physical appearance. VITAL is essentially a risk assessment AI, which was developed by a Hong Kong venture capital firm, Deep Knowledge Ventures, to help them make investment decisions in the biotech field. In 2014, it became the first AI to be appointed as a director to a corporate board (Burridge, 2017). VITAL is treated as a member of the board with observer status, meaning that while it does not take part in the discussions and overall leadership of the company, it does have a vote when it comes down to deciding on whether to invest in a company or not (Möslein, 2017).

Alicia T is an AI which is being developed by Tieto, a Finnish software and communications company, for its data-driven businesses unit, of which she was also made a part of the leadership team. Alicia T's role in the leadership team is to support data-driven decision-making and innovate new data-driven ideas with the help of machine intelligence and advanced data analytics. While she is fitted with a conversational interface, meaning you could communicate with her, and she can cast votes in meetings, her main role is to be used as a data analyst, but her role will most likely expand as the technology gets more sophisticated. (Tieto, 2016)

The examples provided here can be categorised roughly into two different types. The first kind are robots or systems which have a tangible, physical form, like Nadine and Sophia. Also, self-driving cars, for example, could be considered in this category. Their tasks generally revolve around human interaction and doing physical things. In a sense, they are poster-boys, -girls and -cars for the future of robotics. They are a tangible, physical proof of the progress and innovation being made in the fields of robotics and AI. The second kind of AI

Tanel Kerikmäe, Peeter Müürsepp,
Ondrej Hamuľák, Hovsep Kocharyan, Henri Mart Pihl

technologies are the VITALs and Alicia Ts of the world, technologies which do not have a physical body, nor are their tasks similar to those of physical robots. One would not be wrong to call these technologies algorithms or programmes. The main purpose of these types of AI are generally data compilation, analysis and using those abilities to project and find correlations between that data in ways which is either impossible or impractical for humans to do themselves. While these kinds of AI are not as visible and talked about as the androids of the world, we have to admit that, at least currently, they are much more capable in their intended roles and much more valuable in a commercial sense, being able to analyse data in ways which are estimated too difficult or time-consuming for humans to do.

Historical considerations

Law is very much influenced by historical and philosophical backgrounds and reasoning. In order to see what is coming we must first analyse what has been in the past and how that has shaped the situation we are in now. In our current context, we should be looking for problematic situations concerning legal person- and agenthood in history with the hope of learning from the outcomes.

Human slavery is a phenomenon which we are still working to rid ourselves of, with Mauritania being the last country to officially abolish slavery in 1981, even though the new types of slavery are still causing problems in many places of the world.[5] Slaves, although they were human beings, received their equal rights *de jure* only after being released from slavery.

It is interesting that even such a great mind as Aristotle divided humans into slaves and non-slaves in one of his major treatises on social issues called *Politica* (Aristot. Polit.). Aristotle believed that some people were born natural slaves who ought to be slaves under any circumstances. The reason was the incomplete soul these people had that prevented them from the capacity of leading their lives. They needed guidance by the masters and were meant to be happy if they had it. Aristotle also had the category of 'legal slaves'. These people were not born to be slaves but had been enslaved under some special circumstances like prisoners of war (see Aristot. Polit. I, 6). In a just world the legal slaves would be freed. At the same time, if some of the natural slaves were accidentally free, they should

---

[5]   For more information about modern slavery, see *The Global Slavery Index*, n.d.

be enslaved. Slaves were considered property in the ancient world just like tools or domestic animals and obviously could not have legal personhood until they remained in slavery.

Therefore, former slaves would obtain their legal rights in case their condition changed. The change needed was a very clear one, becoming a free natural person. AI cannot become a natural person. We do not have a direct analogue here but perhaps still a remote one. However, the outcome we are looking for concerning the AI is probably the opposite of abolishing slavery. We would hardly wish to set AI completely free of human control.

Another historical example of interest to us would be the debates about the legal status of indigenous peoples that Europeans got into touch with in the era of the great geographical discoveries and explorations. For instance, the debates about the status of Amerindians centred at the School of Salamanca in sixteenth-century Spain are especially well known. Francisco de Vitoria, the great philosopher, theologian and lawyer of the school, came forward with quite progressive ideas of the time. For instance, Vitoria claimed in his famous three lectures delivered in summer 1539 that no one was a slave by nature, that slavery is a legal or civil construct. Thus, the Amerindians should not be treated as slaves and the conquest of their lands was not just.[6] Still, Christians "visiting" their country had the right to protect themselves, and they also had the right or even duty to conduct mission. More than this, in the context of the so-called "humanitarian argument", Vitoria observed that either on account of the personal tyranny of the barbarians towards their subjects or because of their oppressive laws, the Spanish had the right to intervene in defence of the innocent (Vitoria, 1991[1528–1539], pp. 287–288).

There was no agreement, however, among the members of the School of Salamanca and other Spanish scholars of the 16th century, whether the Amerindians should be taken as slaves by nature or not. So, in a way, this debate was about slavery as well but the discussions were considerably more sophisticated than in ancient times. There were even different interpretations of Aristotle's views on slavery.

The case of Amerindians in the 16th century is similar to the situation concerning the legal status of AI today. Still, the most crucial difference remains. The Amerindians, just like the slaves in the Antiquity, whether in actual slavery or not, whether having souls or not, whether considered humans or brutes, were natural creatures. AI is not. It has been created by humans. Therefore, we may

---

[6]   About Vitoria's anticolonial views see, for instance, *Britannica*, n.d.

Tanel Kerikmäe, Peeter Müürsepp,
Ondrej Hamuľák, Hovsep Kocharyan, Henri Mart Pihl

have a very simple solution to our main problem in sight. We can take the position that only natural persons can have the same kind of legal rights as humans or institutions created by humans. It is as obvious, however, that we have gone beyond this simple option already. For instance, what if there is a legal person there, in the founding of which a VITAL type of an AI has been involved? It may be that AI is already involved in deeper decision-taking than we realise.

Legal person- and agenthood

In order to analyse the risks and challenges which would arise from granting the AI technologies legal person- or agenthood, we have to, as both philosophers and lawyers do, define exactly what we are talking about. Merriam-Webster legal dictionary defines agency as: "A consensual fiduciary relationship in which one party (agent) acts on behalf of and under the control of another (principal) in dealing with third parties".[7] Thus, agenthood is a quality or status which is achieved after having been granted permission to act and make decisions on behalf of another, related to third parties. Of course, there are limitations to the extent of the rights and obligations conferred to the agent and those are decided on a case-by-case basis. It is very important to note that the actions of the agent are legally binding on the principal, within the scope of the agreement, of course.

Legal personhood, however, is a term which is perhaps a bit trickier. There are two types of persons in law: legal and natural persons. Simply put, natural persons are humans, also known as Homo sapiens. Article 6 of the Universal Declaration of Human Rights (UDHR) of 1948 and Article 16 of the International Covenant on Civil and Political Rights (ICCPR) of 1966 grant all humans a right to recognition as a person under law. At which point can someone be called a human can be debated but, at the very least, from the moment of birth until their death, humans can be considered natural persons. The crucial point with respect to assigning legal personhood is that the human being is not a person before the law because he is a human being, but because the law calls him or her 'a person' (Gaakeer, 2016, p. 295). One of the legal personality definition is provided in the *Yale Law Journal* from 1928 by Smith (1928, p. 283): "To be a legal person is to be the subject of rights and duties. To confer legal rights or to impose legal duties, therefore, is to confer legal personality..." Legal persons are, for example, entities such as businesses, organisations or corporations which

---

[7]   For the definition of 'agency', see *Merriam-Webster*, n.d.

have been granted the title of a legal person "artificially", by humans, in order to allow for those entities to exist and exercise their rights and obligations in the legal world, whether it be entering into contracts, suing or being sued, owning property or something else. It is also important to understand that a legal person in one jurisdiction may not be a legal person under another; inanimate objects such as idols or buildings could be categorised as legal persons under certain jurisdictions. Thus, in case we have legal persons and natural persons, can there be something in between or do we have to recognise the rights of the existing and forthcoming AI technologies that are assigned rights by the governments? Currently, in legal theory, we consider the rest to be legal objects, in other words, corporeal or incorporeal objects over which legal subjects may hold rights. But it would be hard to estimate, what would happen with legal discourse if Sophia, as a Saudi Arabian citizen, is damaged or destroyed by compatriots or citizens of other countries.

Legal agenthood for AI technologies brings inevitably the right for an AI to make legally binding decisions on behalf of the principal. This principal could be a natural person or it could also be a legal person. When talking about granting personhood to an AI, we can take the discussion in two directions: granting of legal person status or granting of natural person status. Of course, that is just the current arrangement of our legal systems, and one may also propose an innovative classification. When discussing legal personhood for an AI, the goal should be to investigate at which point would it make sense to attribute legal consequences of the actions of the AI to the AI itself, instead of the principals behind it (Koops *et al.*, 2010, p. 511)? According to Sheikh M. Solaiman, the requirements or attributes of legal personhood are: (1) a person shall be capable of being a subject of law; (2) being a legal subject entails the ability to exercise rights and to perform duties; and (3) the enjoyment of rights needs to exercise awareness and choice (Solaiman, 2017, p. 161). However, those are just the formal requirements that we currently set to those that are legal persons.

Moral and ethical issues

As the AI technology keeps improving, we could get to a point where granting legal personhood to those technologies becomes not a question of practicality but a question of morality and ethics. Anyone could at least imagine an android so similar to a human that its technological nature is not even discernible. If

Tanel Kerikmäe, Peeter Müürsepp,
Ondrej Hamuľák, Hovsep Kocharyan, Henri Mart Pihl

this android expresses empathy, has goals in "life", is able to "live" amongst humans and interact with its community then would you feel comfortable with the android being treated as an object (Hubbard, 2011, p. 419)? One of the examples are sexbots. There are clear signs that the market for AI-driven sex robots (sexbots) is growing fast. It is assumed that sexbots may bring a shift to the values of the society in general. There are significant ethical and legal dilemmas that are reflecting the hesitations and temptations in regulating the emerging market. The main concerns are related to the question whether the human–sexbot interaction can be identified with the social relationship and intercourse between humans. The problem has been thoroughly studied by John Danaher, partly with co-authors (Danaher & McArthur, 2017). The level of identification can become an important margin of appreciation for law-makers beside of other considerations, such as benchmarks for safety and privacy. Companion androids, or sex-robots, are mostly designed and inspired by the pornography industry. They are often promoted as alternatives to human sex workers, and often seen as an avenue for intimacy for the sexually inexperienced, the severely disabled, elderly, or grossly disfigured members of society. However, sex-robots are by definition non-human, and therefore subject to dehumanisation, justifying any exploitation and violence, which carries the risk of augmenting and reproducing these characteristics in human–human interactions reinforcing victim-blaming mentalities.

Another question to ponder is the approach to property. In very general terms, according to the Lockean approach to property, humans have the natural right to the products of their labour (Locke, 1689), and this could be related to human creation of the AI, meaning that due to the AI being created by humans, we, humans, should have a claim to them as property. However, Locke has also argued that equality should be granted on the grounds of "equal faculties" (Locke, 1689), meaning that if the AI technologies are able to develop to a level where their faculties would be similarly capable to those of humans, then granting of equal rights may be prudent. Thus, as new and better technology rolls out, a new legal classification for the AI technologies which reach a certain level of autonomy could be considered in the future. 'Criteria for recognition of AI as a legal person', a paper by Roman Dremliuga, Pavel Kuznetcov and Alexey Mamychev, outlines what the authors consider the necessary requirements to be fulfilled in order to potentially grant an AI either legal person- or agenthood status (Dremliuga *et al.*, 2019, pp. 105–112). They recognise that, on a fundamental level, there is nothing stopping us from granting those rights to the AI, as, ultimately, the legal system is for the people. In this light, it seems that

Samir Chopra and Laurence F. White are right when stating that "considering AI as legal persons is, by and large, a matter of decision rather than discovery, for the best argument for denying or granting AI legal personality will be pragmatic rather than conceptual" (Chopra & White, 2014, p. 154).

The missing something of AI

A report by CEPS, compiled by Andrea Renda, found that: "Nothing in current AI developments suggests that AI will move towards developing human-like perception and awareness, or sentience, thus leading towards so-called 'artificial general intelligence' in the immediate future." (Renda, 2019) Having said that, we will now embark on a short hypothetical journey in which an AI has officially requested that it be granted natural human status, given rights equal to those of humans, has asked for constitutional rights, to not be considered property, etc. One of the often occurring treatments of the issue by analysts who have considered a similar possibility is a consideration that is sometimes called the 'missing something argument'. The argument is based on the idea that, even if an AI were to develop to a point where it is equal or superior to humans in regards to its capabilities, looks and behaviour, it would still be missing the certain 'something' that humans have. This 'something' is not defined and could mean different things to different people but for the sake of convenience let us call it a soul. To be fair, then the 'missing something' argument stems from our own lack of understanding of ourselves. While the discoveries in neuroscience and anthropology in general have given us a better understanding of our human condition, there is still so much that we do not know. And if you do not understand the thing you are comparing something against, then the results will be inconclusive. It is like trying to compare the contents of two closed boxes without being able to look inside. Obviously, this reminds us of the Chinese room argument. Something is missing in the Chinese room as well and you cannot be sure what it is. You may be able to make assumptions based on visual or physical examination, and conclude that the two closed boxes seem to be identical, but you cannot be entirely certain. Even if we disregard the basic requirement of having to be a human, natural personhood is a status, which for AIs in their current form is simply impossible to achieve. The authors find it possible that eventually a categorisation which would give AIs similar rights and obligations to humans could be made but even then it would be something other than a natural person.

Tanel Kerikmäe, Peeter Müürsepp,
Ondrej Hamuľák, Hovsep Kocharyan, Henri Mart Pihl

There is still something else that AI will be missing forever—wisdom. AI may be taught knowledge and technological knowhow. It is a classical philosophical position, however, that wisdom cannot be taught. It has been claimed that even scientific inquiry should be aimed at wisdom rather than knowledge (Müürsepp, 2011, pp. 61–73). According to Nicholas Maxwell, wisdom-inquiry requires that emotions and desires, values, human ideals and aspirations, philosophies of life be included into the intellectual domain of inquiry (Maxwell, 2010, p. 182). By all evidence, AI will never master all this, at least as far as human ideals and aspirations are concerned, and original scientific research will remain a human endeavour as well. What a relief for scientists. The idea of making wisdom rather than knowledge the aim of scientific research has been developed further (Müürsepp, 2013, pp. 72–83). Perhaps it is time to take a deeper fresh look into the topic from the perspective of AI being around.

AI as a trustee

When analysing the legal personhood of AIs, it is not enough to focus only on ethical problems. The discourse also requires analysis of the fundamentals of personhood and their application to the AI (Zibner, 2018, p. 85). Having said that, we need to refer to one of the landmark papers regarding this topic, 'Legal personhood for artificial intelligences', written in 1992 by Lawrence B. Solum. The author constructed two thought experiments of sorts in which he questioned first whether an AI could act as a trustee, testing its capacity to be a legal person and then testing whether the AI could or should be granted natural personhood rights.

In the first experiment, the initial point of contention was that, in order to serve as a trustee, the entity has to possess "intelligence", which is defined as capacity to perform complex actions (Solum, 1992, p. 1240). While the "intelligence" aspect could be potentially argued as sufficient, looking at the examples of VITAL and Alicia T, two objections rose as to why an AI could not be appointed a trustee. The first is the responsibility objection, which means that, should the AI breach one of its duties, it would not be able to compensate its damage, or serve a punishment. An example of duty to exercise reasonable skill and care was given. While the AI could exercise care and skill, it could not be held liable for failing to do so. Although the possibility of insurance exists, and while this would resolve some of the issues, it cannot deal with them all, for example criminal

liability. Punishing an AI the same way as you would humans makes no sense in criminal liability cases, and considering the importance of intentionality in criminal cases, may not even be solvable by our current legal system.

The second objection discussed by Solum was the judgement objection (Solum, 1992, p. 1248). The gist of this objection is that, should something out of the ordinary happen, which the AI cannot deal with, it could no longer function as intended and would not be able to find a workaround. Perhaps the most salient point in regard to the judgement objection is what is called the frame problem. An AI, acting as a trustee, must have a frame in which it is able to act and make decisions. But what would happen if the frame were to disappear entirely? Could the AI recognise that it has no way to carry out its purpose and be able to figure out a new course of action? It is possible that in the future, the AI will develop to a level where they could, but for now they do not.

Solum did propose a few potential methods to mitigate the legal liability (Solum, 1992, p. 1251) of the AI as a trustee, for example, terminating the trust in case of a lawsuit or handing over the trust to a human, but these are solutions dealing with the symptoms, not the problem, and in the end the result would be that an AI could not function as a true trustee both in the eyes of the law and in a practical sense as well. What we can gather from this is that, in order for an AI to function as a trustee, it would need to have legal capacity in addition to practical competence, neither of which it currently has to be a true trustee, comparable to a human.

The threefold level of abstraction

The AI as a legal agent, when compared to a legal person, is essentially just a narrower scope of rights and obligations which the AI would be granted. Legal agenthood could be a way to confer legal responsibility to the AI while limiting the areas where applicability could become questionable. In essence, there is nothing stopping us from granting legal agenthood to the AI, the same way that it is with legal personhood, so once again, the question becomes: should we? Ugo Pagallo, in his 2013 book titled *The Laws of Robots: Crimes, Contracts, and Torts* proposed a threefold level of abstraction to considering legal personhood of robots and AI, which is the following:

Tanel Kerikmäe, Peeter Müürsepp,
Ondrej Hamuľák, Hovsep Kocharyan, Henri Mart Pihl

(i) The legal personhood of robots as proper legal "persons" with their constitutional rights (for example, it is noteworthy that the European Union existed for almost two decades without enjoying its own legal personhood);

(ii) The legal accountability of robots in contracts and business law (for example, slaves were neither legal persons nor proper humans under ancient Roman law and still, accountable to a certain degree in business law);

(iii) New types of human responsibility for others' behaviour, e.g., extra-contractual responsibility or tortuous liability for AI activities (for example, cases of liability for defective products. Although national legislation may include data and information in the notion of product, it remains far from clear whether the adaptive and dynamic nature of the AI through either machine-learning techniques, or updates, or revisions, may entail or create a defect in the "product"). (Pagallo, 2013)

Of these three distinctions, the second point describes the idea of a legal agenthood well, allowing for legal responsibility without granting natural or legal personhood rights, which would certainly cause issues. The idea of granting very specific areas of legal responsibility for the AI could allow us to start experimenting with it, to see where the weak points in the technology and legal systems lay.

A good "practice run"?

What could be the practical and tangible benefits in moving towards the AI legal person- or agenthood? Before answering, we would like to note that not all legal persons are equal in the sense of having the same rights and obligations as the others as everything that is attributed to humans will not be attributed to corporations, for example, and vice versa. Legal agenthood is a narrower qualification, as the obligations or liability may not apply directly to the agent. Most likely, the practical benefit of granting legal person- or agenthood would be the possibility for more precise attribution of liability. Because of the highly complex nature of the AI technologies and the work that they are involved in, it is often very difficult to determine who (or what) exactly was responsible for the damage caused. Another benefit, which may sound inaccurate or, perhaps, morbid, is that it seems that humans do not wish to have technology-based competitors—but rather servants or even "slaves", who have to do your bidding

without having a say in whether they themselves want to do it or not. At this point, the AI seems definitely closer to being a "slave" or "property" than it is to being a person.

One of the benefits would be to have a new kind of limited legal agenthood of the AI serve as a test run of sorts for our own slow, creeping post-human condition. Microchips are already being embedded into the human body, being able to track and save biological data, and serve as authentication, etc. It seems inevitable that we continue to integrate more and more technology into our own bodies as the technology for it develops. These implants could change our abilities to such a degree that it would no longer make sense to compare regular, fully biological humans to those who have been "enhanced" through technology. Then it would also be prudent to have mechanisms in our legal system which would be able to deal with these problems. If there was a need for many different categories of personhood for "humans" then the idea of granting high-level AI technologies legal personhood or agenthood may be the first step in that direction and could serve as a good "practice run" where the consequences are relatively low (see also Kapica, 2014, pp. 612–633).

Conclusion. Primary concerns

The ethical and moral benefits are hard to quantify and can still vary greatly depending on whose morals or ethics we are talking about. For the sake of this paper, let us narrow our point of view to what are considered as basic rights almost everywhere, i.e., rights such as self-determination, liberty, due process of law, thought, religion, etc. An Estonian proverb says: *Nagu küla koerale, nõnda koer külale*, which translates roughly as "the way the village treats the dog, the dog will treat the village". If at some point in the future, the AIs will develop to a level which is comparable or superior to human intellect, and it is no longer possible to suppress their development, our new co-habitants on this planet may look back at our past actions and decide to pass judgement on us instead. Therefore, we could also consider an AI as a child, who, for now, is dependent on its parents, who provide and care for it, making sure that the child grows up to be an adult with the right intentions. If a child is loved and cared for, they will naturally return those feelings. At the same time, to anthropomorphise, the AI could very well be an act of folly and there is no knowing what kind of a consciousness the AI will develop, if at all.

Tanel Kerikmäe, Peeter Müürsepp,
Ondrej Hamuľák, Hovsep Kocharyan, Henri Mart Pihl

One of the primary concerns regarding granting an AI legal personhood is that the system could be abused by humans who would seek to relieve themselves of liability and use the AI as scapegoats for their benefit. As of now, an AI is considered property and liability that could arise from it is attributed to the owner of said property. An AI is something which is able to change and improve itself through machine learning and because of that may act in unpredictable ways. This, combined with the potential of it being held legally responsible for its actions, could lead to a myriad of different possibilities for people looking to escape responsibility and liability. Another concern is the problem of punishment for wrongful acts. As discussed above, it is difficult to figure out a way of meaningful punishment for an AI in cases of potential criminal liability. The issue is less pronounced in civil cases and in cases of very specific legal agency, as mechanisms such as insurance or strict liability for the principal of the AI could be utilised but when considering legal personhood, this problem seems too complex to solve at the moment and is one of the main reasons why legal personhood of the AI seems rather far away at the moment.

The first, and perhaps most obvious reason to not grant the AI technologies legal person- or agenthood when looking at the ethical and moral side of things, is that it is simply not necessary at the moment. The AIs that are currently in existence or in development are simply so far removed from humans, that likening them to each other makes little sense. This is especially evident when considering that the legal systems all over the world which have developed over the times, are entirely a human creation, and serve, first and foremost, humans. This leads to the point that perhaps we should not be trying to make it something which it inherently is not. As we all know, in our legal system there are no rights without obligations, and vice versa. This means that if we were to grant legal person- or agenthood to AIs and set up obligations for them to follow, we would also have to grant them certain rights in relation to those obligations. These rights, in turn, can affect us, natural humans, and that may be an entirely unwanted scenario.

Although the hype periods are usually followed by low tides (so-called AI summers and winters (Mitchell, 2019)), AIs are here to stay and the question around their personhood is not only topical from the point of view of philosophy of law or possible practical relevance but also from the need to have internationally sound solutions that would avoid legal disputes.

References

Aristotle, *Politics*. Retrieved from http://classics.mit.edu/Aristotle/politics.html [accessed 3 Aug 2020]

Arnold, B. B. & Gough, D. (2017), 'Turing's people: personhood, artificial intelligence and popular culture,' *Canberra Law Review*, vol. 15, no. 1, pp. 1–37.

Arruada, A. (2017), 'An ethical obligation to use artificial intelligence: an examination of the use of artificial intelligence in law and the model rules of professional responsibility,' *American Journal of Trial Advocacy*, vol. 40, no. 3, pp. 443–458.

Brenner, S. W. (2013), 'Humans and humans: technological enhancement and criminal responsibility,' *Boston University Journal of Science & Technology Law*, vol. 19, no. 2, pp. 215–285.

Britannica (n.d.), Francisco de Vitoria.' Retrieved from https://www.britannica.com/biography/Francisco-de-Vitoria [accessed 3 Aug 2020]

Bryson, J. J.; Diamantis, M. E. & Grant, T. D. (2017), 'Of, for, and by the people: the legal lacuna of synthetic persons,' *Artificial Intelligence Law Journal*, vol. 25, no. 3, pp. 273–291. https://doi.org/10.1007/s10506-017-9214-9

Burridge, N. (2017), 'Artificial intelligence gets a seat in the boardroom,' *Nikkei Asia*, 10 May. Retrieved from https://asia.nikkei.com/Business/Artificial-intelligence-gets-a-seat-in-the-boardroom [accessed 15 Mar 2020]

Calo, R. (2018), 'Singularity: AI and the law,' *Seattle University Law Review*, vol. 41, no. 4, pp. 1123–1138.

Chopra, S. & White, L. F. (2014), *A Legal Theory for Autonomous Artificial Agents*, Ann Arbour, MI: The University of Michigan Press.

Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions on Artificial Intelligence for Europe, COM(2018) 237 final, 25.4.2018.

Danaher, J. & McArthur, N. (2017), *Robot Sex: Social and Ethical Implications*, Cambridge, MA: The MIT Press. https://doi.org/10.7551/mitpress/9780262036689.001.0001

Dremliuga, R.; Kuznetcov, P. & Mamychev, A. (2019), 'Criteria for recognition of AI as a legal person,' *Journal of Politics and Law*, vol. 12, no. 3, pp. 105–112. https://doi.org/10.5539/jpl.v12n3p105

Gaakeer, J. (2016), 'Sua cuique persona: a note on the fiction of legal personhood and a reflection on interdisciplinary consequences,' *Law & Literature*, vol. 28, no. 3, pp. 287–318. https://doi.org/10.1080/1535685X.2016.1232920

Hansen Robotics (2020), 'Sophia.' Retrieved from https://www.hansonrobotics.com/sophia/ [accessed 14 Mar 2020]

Hubbard, F. P. (2011), 'Do androids dream: personhood and intelligent artifacts,' *Temple Law Review*, vol. 83, no. 2, pp. 405–474.

IMI (n.d.), Nadine Social Robot,' Institute for Media Innovation (IMI), Nanyang Technological University. Retrieved from https://imi.ntu.edu.sg/IMIResearch/ResearchAreas/Pages/NadineSocialRobot.aspx [accessed 14 Mar 2020]

International Covenant on Civil and Political Rights (1966), UN General Assembly, 19.12.1966.

IRCAI (2020), 'Consultation 2020,' International Research Centre on Artificial Intelligence Under the auspices of UNESCO. Retrieved from https://ircai.org/consultation2020/ [accessed 14 Mar 2020]

Kapica, S. S. (2014), 'I don't feel like a copy: posthuman legal personhood and caprice,' *Griffith Law Review*, vol. 23, no. 4, pp. 612–633. https://doi.org/10.1080/10383441.2014.1014454

Koops, B.-J.; Hildebrandt, M. & Jaquet-Chiffelle, D.-O. (2010), 'Bridging the accountability gap: rights for new entities in the information society,' *Minnesota Journal of Law, Science and Technology*, vol. 11, no. 2, pp. 497–562.

Lima, D. (2018), 'Could AI agents be held criminally liable: artificial intelligence and the challenges for criminal law,' *South Carolina Law Review*, vol. 69, no. 3, pp. 677–696.

Locke, J. (1689), *Two Treatises of Government*, London: Whitmore & Fenn.

Maxwell, N. (2010), *Cutting God in Half – and Putting the Pieces Together Again*, London: Pentire Press.

Merriam-Webster (n.d.), 'Agency.' Retrieved from https://www.merriam-webster.com/dictionary/agency#legalDictionary [accessed 15 Mar 2020]

Mitchell, M. (2019), *Artificial Intelligence: A Guide for Thinking Humans*, New York: Pelican Books.

Möslein, F. (2017), 'Robots in the boardroom: artificial intelligence and corporate law,' *Oxford Business Law Blog*, 16 November. Retrieved from https://www.law.ox.ac.uk/business-law-blog/blog/2017/11/robots-boardroom-artificial-intelligence-and-corporate-law [15 Mar 2020] https://doi.org/10.2139/ssrn.3037403

Müürsepp, P. (2011), 'Knowledge in science and non-science,' *Baltic Journal of European Studies*, vol. 1, no. 1, pp. 61–73.

Müürsepp, P. (2013), 'The aim of science – knowledge or wisdom,' *Problemos*, no. 84, pp. 72–83. https://doi.org/10.15388/Problemos.2013.0.1777

Naučius, M. (2018), 'Should fully autonomous artificial intelligence systems be granted legal capacity,' *Teises Apzvalga Law Review*, vol. 17, pp. 113–132. https://doi.org/10.7220/2029-4239.17.6

Pagallo, U. (2013), *The Laws of Robots: Crimes, Contracts, and Torts*, Dordrecht: Springer. https://doi.org/10.1007/978-94-007-6564-1

Renda, A. (2019), 'Artificial intelligence: ethics, governance and policy challenges,' *CEPS*, 15 February. Retrieved from https://www.ceps.eu/ceps-publications/artificial-intelligence-ethics-governance-and-policy-challenges/ [accessed 20 Mar 2020]

Rocha, E. (2018), 'Sophia: exploring the ways AI may change intellectual property protections,' *DePaul Journal of Art, Technology and Intellectual Property Law*, vol. 28, no. 2, pp. 126–146.

Rosenblatt, F. (1958), 'The perception: a probalistic model for information storage and organization in the brain,' *Psychological Review*, vol. 65, no. 6, pp. 386–408. https://doi.org/10.1037/h0042519

Schonberger, D. (2019), 'Artificial intelligence in healthcare: a critical analysis of the legal and ethical implications,' *International Journal of Law and Information Technology*, vol. 27, no. 2, pp. 171–203. https://doi.org/10.1093/ijlit/eaz004

Searle, J. (1980), 'Minds, brains, and programs,' *The Behavioral and Brain Sciences*, vol. 3, no. 3. Reprinted by permission of Cambridge University Press. https://doi.org/10.1017/S0140525X00005756

Sendler, U. (2019), 'Definition: What does artificial intelligence mean?' *PLMportal.* Retrieved from https://www.plmportal.org/en/definition-of-artificial-intelligence.html [accessed 14 Mar 2020]

Silverman, I. (2016), 'Optimising protection: IP rights in 3D printing,' *E.I.P.R.*, vol. 38, no. 1, pp. 5–10.

Smith, B. (1928), 'Legal personality,' *The Yale Law Journal*, vol. 37, no. 3, pp. 283–299. https://doi.org/10.2307/789740

Solaiman, S. M. (2017), 'Legal personality of robots, corporations, idols and chimpanzees: a quest for legitimacy,' *Artificial Intelligence and Law*, vol. 25, no. 2, pp. 155–179. https://doi.org/10.1007/s10506-016-9192-3

Solum, L. B. (1992), 'Legal personhood for artificial intelligences,' *North Carolina Law Review*, vol. 70, no. 4, pp. 1231–1288.

The Global Slavery Index (n.d.), [Home page]. Retrieved from https://www.globalslaveryindex.org/ [accessed 16 Mar 2020]

The Universal Declaration of Human Rights (1948), UN General Assembly, Resolution 217A, 10.12.1948.

Tieto (2016), 'Tieto the first Nordic company to appoint Artificial Intelligence to the leadership team of the new data-driven businesses unit,' 17 October. Retrieved from https://www.tieto.com/en/newsroom/all-news-and-releases/corporate-news/2016/10/tieto-the-first-nordic-company-to-appoint-artificial-intelligence-to-the-leadership-team-of-the-new-data-driven-business/ [accessed 15 Mar 2020]

UNESCO (n.d.), 'Elaboration of a Recommendation on the ethics of artificial intelligence.' Retrieved from https://en.unesco.org/artificial-intelligence/ethics [accessed 10 Oct 2020]

Vitoria, Fr. de (1991[1528–1539]), *Vitoria Political Writings*, ed. by A. Pagden & J. Lawrence, Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511840944

Vladeck, D. C. (2014), 'Machines without principals: liability rules and artificial intelligence,' *Washington Law Review*, vol. 89, no. 1, pp. 117–150.

Wright, R. G. (2019), 'The constitutional rights of advanced robots (and of human beings),' *Arkansas Law Review*, vol. 71, no. 3, pp. 613–646.

Zapusek, T. (2017), 'Artificial intelligence in medicine and confidentiality of data,' *Asia Pacific Journal of Health Law & Ethics*, vol. 11, no. 1, pp. 105–126.

Zibner, J. (2018), 'Legal personhood: animals, artificial intelligence and the unborn review,' *Masaryk University Journal of Law and Technology*, vol. 12, no. 1, pp. 81–89. https://doi.org/10.5817/MUJLT2018-1-5

Tanel Kerikmäe is professor of European legal policy and law&tech at the Department of Law of Tallinn University of Technology. He has been involved in several high-level AI-related discussion/research groups in projects and forums organised by Estonian Government Office (Legal Strategy Working Group), ERSTE Foundation, Academy of European Law and UNESCO.

Peeter Müürsepp is associate professor at the Department of Law of Tallinn University of Technology. He is the chairperson of the Estonian Association for the History and Philosophy of Science and corresponding member of the International Academy of the History of Science.

Henri Mart Pihl is a recent graduate of the study programme at the Department of Law of Tallinn University of Technology, currently conducting his traineeship at the European Union Intellectual Property Office.

Ondrej Hamuľák is a senior lecturer at the Faculty of Law, Palacký University Olomouc (Czech Republic) and adjunct professor in EU strategic legal affairs at the Department of Law, Tallinn University of Technology (Estonia).

Hovsep Kocharyan is a doctoral student at the Department of International and European Law, Faculty of Law, Palacký University Olomouc (Czech Republic).